

Stretching force dependent transitions in single stranded DNA

Kulveer Singh¹, Surya K. Ghosh¹, Sanjay Kumar², and Anirban Sain^{1*}

¹ Department of Physics, Indian Institute of Technology, Bombay, Powai, Mumbai-400 076, India.

² Department of Physics, Banaras Hindu University, Varanasi 221 005, India

(Dated: March 1, 2013)

Mechanical properties of DNA, in particular their stretch dependent extension and their loop formation characteristics, have been recognized as an effective probe for understanding the possible biochemical role played by them in a living cell. Single stranded DNA (ssDNA), which, till recently was presumed to be an simple flexible polymer continues to spring surprises. Synthetic ssDNA, like polydA (polydeoxyadenosines) has revealed an intriguing force-extension (FX) behavior exhibiting two plateaus, absent in polydT (polydeoxythymidines) for example. Loop closing time in polydA had also been found to scale exponentially with inverse temperature, unexpected from generic models of homopolymers. Here we present a new model for polydA which incorporates both a helix-coil transition and a over-stretching transition, accounting for the two plateaus. Using transfer matrix calculation and Monte-Carlo simulation we show that the model reproduces different sets of experimental observations, quantitatively. It also predicts interesting reentrant behavior in the temperature-extension characteristics of polydA, which is yet to be verified experimentally.

PACS numbers: 87.10.Pq, 87.15.La, 05.70.Fh

In order to understand how the mechanical properties of DNA and RNA influence biological processes like transcription and translation in living cells and viruses, these biopolymers are stretched in vitro to study their nonlinear elasticity and their internal structure. ssDNA, despite receiving less attention compared to double stranded DNA (dsDNA) [1, 2], has recently attracted a lot of interest. Smith et al. [3] showed that the FX diagram for wild type λ -phase ssDNA can be described by FJC model only at low force ($< 10pN$). To explain the behavior at higher force they used a *modified* freely jointed chain (mFJC) model [1, 3] with stretch dependent kuhn length. Subsequently, it was discovered that synthetic ssDNA have interesting sequence dependent properties. For example, polydA was found to have a higher bending rigidity than polydT and as a consequence polydA takes relatively longer time to form a loop [4]. Further, the loop formation time varies exponentially with inverse temperature which can neither be explained by FJC nor worm like chain (WLC) model, which are generic models for flexible and semiflexible polymers, respectively. This behavior was attributed to strong stacking interaction among pyrimidine bases in polydA [4–6]. Subsequently, FX characteristics of polydA revealed that it undergoes two successive transitions under external stretching [7, 8] generating two distinct plateaus in the force extension (FX) curve. The first one at $\sim 23pN$ force was proposed to be a helix to coil transition in which the inter-base stackings are broken and helical polydA segments transform to a polydA coil. The second transition at $\sim 114pN$ force was attributed to the over-stretching of the constituent bases. It was conjectured [3, 7] that over-stretching results from the conformational change of the sugar molecules from C3'-endo to C2'-endo pucker conformation.

Zimm-Bragg model [9] which was originally proposed to explain temperature driven helix-coil transition in proteins, has been used to study the force driven helix-coil transition in polydA. Theoretical models have used mean field approximation [11] and exact evaluation of partition function [12] to

explain the first plateau experimentally seen [10] in the FX diagram of polydA at low forces ($< 60pN$). But the second plateau, involving the over-stretching transition is beyond the scope of these models. Overstretched in dsDNA has been studied using Ising like two state models by various groups [13, 14]. Here we propose a model which quantitatively reproduce both the force driven behavior as well as zero force conformational fluctuations like the loop formation time, as observed experimentally. Double plateau behavior has also been addressed before [15, 16], theoretically, albeit using a lattice model. But lattice models generically underestimate entropic effects and also quantitative comparisons to experiments were not possible.

Model: We model polydA as a chain of connected segments, each representing a nucleotide. Length of a segment (bond) represent the phosphate to phosphate distance in the ssDNA backbone. The bond length can be l_h , l_c or l_s depending on whether the segment is in the helix ($l_h = 0.37nm$ [10]), coil ($l_c = 0.59nm$ [2]), or overstretched-coil state ($l_s = 0.7nm$ [7]). While the value of l_c has been most well documented [2] in the literature, l_h is determined [10] by noticing the 1.6 times increase in contour length of polydA upon helix to coil transition. Here l_h of course is the projected length of the helical contour on the central axis of the helix. Further, $l_s = 0.7nm$ has been inferred from the maximum extension reached by the polydA chain under very high force ($\sim 600pN$) in Ref[7, 8]. Incidentally, $l_s = 0.7nm$ also matches with the distance between the consecutive phosphates when the deoxyribofuranose ring is in the C2'-endo pucker conformation. In our model, state of a segment is characterized by (μ, S) , where μ can take values 0 or 1, corresponding to coil or helix state, respectively. Once in the coil state $\mu = 0$, there are two possible states: the normal coil or the overstretched coil, corresponding to bonds lengths l_c and l_s . These two states are represented by $S = -1$ and $S = +1$, respectively. Hence, there are only three possible states $(1, -1)$, $(0, -1)$ and $(0, 1)$, corresponding to helix, coil

and overstretched-coil. The Hamiltonian of this three state model of polydA (omitting the external force) is

$$H_0 = \sum_{i=1}^N \left(2(1 - \mu_i)\mu_{i+1}\Delta w + \mu_{i+1}\Delta f - JS_iS_{i+1}(1 - \mu_i)(1 - \mu_{i+1}) + hS_i(1 - \mu_i) \right) \quad (1)$$

This Hamiltonian incorporates two transitions, actually cross-overs, since the model is one dimensional. First part of the Hamiltonian involving Δw and Δf , is the simpler version of the original Zimm-Bragg model [9], forwarded by the authors themselves and later used by Tamashiro et al. [12] in the context of forced DNA. In this model the necessary hydrogen bonding, that is required for the formation of helical domain, takes place between adjacent bases (segments), instead of the $i - th$ and $(i + 4) - th$ bases as in the original Zimm-Bragg model [9]. The weight of the configurations cc , hc , ch , hh , in this model, are given by 1, 1, σs and s respectively. Here $\sigma = e^{-2\Delta w/k_B T}$ and $s = e^{-\Delta f/k_B T}$, where $2\Delta w$ is the interfacial energy between the helical and the coil domains and Δf is the difference of energy between the helix and the coil state. The asymmetry between ch and hc arises because a segment (its oxygen atom) can engage either its right neighbor (its hydrogen atom) or its left neighbor to make a hydrogen bond (for details see Ref[9]). Second part of the Hamiltonian is an Ising Hamiltonian, which will be invoked, when the segments are in the coiled or overstretched coil state i.e., when $\mu = 0$. J measures the correlation energy between an adjacent coil and overstretched coil states. $2h$ is the energy difference between a coil and a overstretched coil state. The parameters σ and J are often called the cooperativity parameters of the respective transitions.

We considered the ssDNA chain to be semi-flexible, in which the helical domains have very large persistence length, where as the coiled and overstretched-coil domains have small persistence length. Discretized worm-like-chain hamiltonian for the system is

$$\beta H_{bend} = \sum_{i=1}^{N-1} \left\{ \frac{a_h}{2} \mu_i \mu_{i+1} + (1 - \mu_i)(1 - \mu_{i+1}) \times \left[\frac{a_c}{8} (1 - S_i)(1 - S_{i+1}) + \frac{a_s}{8} (1 + S_i)(1 + S_{i+1}) \right] \right\} \times (1 - \cos \theta_{i,i+1}). \quad (2)$$

Here, $\theta_{i,i+1}$ is the angle between the bond vectors \vec{t}_i and \vec{t}_{i+1} , where $\vec{t}_i = \mathbf{R}_{i+1} - \mathbf{R}_i$ and \mathbf{R}_i is position vector of $i - th$ monomer. The hamiltonian has been constructed in a way such that different bending rigidities are associated when neighboring bonds are of identical type, i.e., hh , cc or ss . It amounts to assuming that helical, coil and stretched-coil domains of the polymer behave like worm-like-chain. a_h , a_c and a_s are the respective persistence lengths. We choose $a_h = 12nm$, $a_c = 1.5nm$ and $a_s = 1.5nm$, i.e., a relatively large persistence length for the helix (still much smaller than that of *dsDNA*, about $50nm$). Theoretical models [10, 12]

assume a_h to be infinite. Our assumption that even ssDNA coils have a small persistence length is in agreement with Seol et al. 's [17] FX data on polyU where they obtained $a_c \sim 1nm$. Also Smith et al [3] reported $a_c \sim 1.5nm$ for wild type λ -phase ssDNA, using mFJC model.

For calculation purpose, we substitute inter-bond angles in terms of bond vectors: $\frac{a}{2}(1 - \cos \theta_{i,i+1}) = \frac{a}{4}(\hat{t}_{i+1} - \hat{t}_i)^2$. Finally, including external stretch \mathbf{F} , the total hamiltonian is

$$H_{total} = H_0 + H_{bend} - \mathbf{F} \cdot (\mathbf{R}_N - \mathbf{R}_0) \quad (3)$$

The force dependent term above can be expressed as $\mathbf{F} \cdot \sum_{i=1}^N l_i \hat{t}_i$ in terms of the bond vectors and their respective lengths l_i , where l_i could be l_h , l_c or l_s depending on the internal state (μ, S) of the bond. Finally, the partition function is $Z(F) = \int \prod_{i=1}^N d\hat{t}_i \sum_{\{\mu_i, S_i\}} \exp[-\beta(H_0 + H_{bend} - FR_x)]$, where $R_x = (\mathbf{R}_N - \mathbf{R}_0) \cdot \hat{x}$ is projection of the end-to-end distance along the force $\mathbf{F} = F\hat{x}$. Using transfer matrix technique we can write $Z(F)$ for the forced chain as,

$$Z = \sum_{\mu_1, S_1, \hat{t}_1; \mu_N, S_N, \hat{t}_N} \langle \mu_1, S_1, \hat{t}_1 | T^{N-1} | \mu_N, S_N, \hat{t}_N \rangle \quad (4)$$

Here the transfer matrix T is an $3mn \times 3mn$ matrix where the internal state space is 3 dimensional, corresponding to the states $(1, -1)$, $(0, -1)$, $(0, 1)$ and the orientation space $\hat{t}_i(\phi, \theta)$ has been discretized into $m \times n$ bins. In Eq.4, $Z(F)$ is obtained as a weighted sum over all the matrix elements of T^{N-1} . We cannot exploit the simplifications normally arising from periodic boundary condition because, here, the polymer has one of its ends fixed and the other end stretched by a force. The details of this calculation can be found in Ref[18]. After calculating $Z(F)$ we can compute $\langle R_x \rangle$ as a function of force using $\langle R_x \rangle = \frac{1}{\beta} \frac{\partial \ln Z}{\partial F}$. This is shown in Fig.1.

Force-extension behavior : In general the stretching force tends to align the chain along the force (\hat{x}), at the cost of entropy. It would also favor the individual bonds to have their highest possible bond lengths i.e., $0.7nm$ in order to maximize the FR_x term. This is achieved only at very high force when the other terms in the Hamiltonian give in to the force term. But at low and intermediate forces the other terms compete. Although helical segments are favored over coil segments (due to s), at low force and at room temperature, entropy has substantial contribution. As a result all the segments are not aligned along the force. That is why the extension per base, at very low force is about $0.2nm$ and not $0.37nm$ (see Fig.1). But bending of the helical domain is disfavored by its large persistence length. Therefore, at low force, the entropy induced bends help some helical segments to convert to coil, which have higher internal energy but lower bending energy compared to helical segments. As the force rises, the FR_x term enforces alignment as well as stretching of the bonds, leading to stacking-unstacking transition which adds extra length (that was previously curled up in helix) to R_x . The abrupt nature of the helix-coil transition results from high cooperativity ($\sigma = e^{-2\Delta w/k_B T} = 0.0015$ here) i.e., the

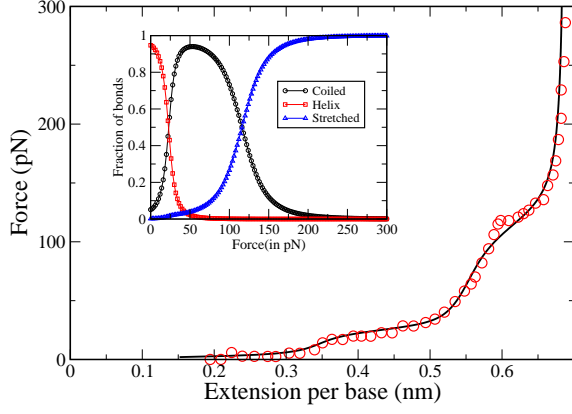


FIG. 1: (color online) Force-Extension curve obtained from our model (solid line), using transfer matrix method, compared with the experimental result of Chen et al. [8] (circle). Such a double plateau feature was first reported by Ke et al and their data is very close to that of Ref[8] shown here. The inset shows how fraction of the different species, namely, helix(circle), coil(square) and stretched coil (triangle) portions change with increasing force, obtained from Monte-Carlo simulation of our model. The intersection points between the helix and coil fractions mark the position of the first plateau ($\sim 23pN$) and that between the coil and overstretched-coil fractions mark the second plateau ($\sim 114pN$) in the FX-plot. Parameters used for our calculation are $2\Delta w = 6.3k_BT$, $\Delta f = -4.93k_BT$, $J = 0.44k_BT$, $h = 1.5k_BT$, where T is the room temperature. We used two different chain lengths, $N = 33$ and 65 , to check N independence of this plot.

coupling term $\mu_i\mu_{i+1}$ in the Hamiltonian. The width of the plateau, i.e., how much length is released at the transition, depends on how many segments are converted from the helical to the coil state. This comes from the abruptness of the jump, shown in the inset. We compute the area under the FX curve: $\int F dx \sim 6k_BT$. Given that $2h = 3k_BT$ is the free-energy change for the second transition, the helix-coil transition accounts for a free-energy change of about $3k_BT$.

Given that polydT and polyU do not have any significant stacking interaction, they might be expected to follow the $\Delta f = \Delta w = 0$ limit of our model. But polydT, polyU also have qualitatively different overstretching behavior than polydA; they overstretch very slowly with force (unlike the transition in polydA). PolyU has been already explained by a modified WLC model [17] and we checked that even polydT can also be explained by the same model but with different parameters.

Loop formation kinetics : Now we discuss the zero force conformational fluctuations resulting from our model. We focus on the observations made by Goddard et al [4] on loop formation properties of short ($N = 8 - 30$ bases) polydA and polydT chains. They had attached complementary base sequences TTGCC and AACGG at the two ends of a polydA/polydT strand and attached fluorophore (F) and quencher (Q) molecules at the ends. This design aimed to detect the formation of hairpin loops by zipping of complemen-

tary base-pairs at the ends. Such a process resulted in quenching of fluorescent intensity of F by Q . They found that, for a given chain length, a polydA chain took longer time than a polydT chain to form a loop. This indicates greater bending stiffness for polydA chains, resulting from stacking interaction between Adenine bases. More intriguing was the result that, loop closing time $\ln(\tau_c) \propto \beta$ for polydA and a nearly flat temperature dependence for polydT.

This process can be approximately described by a two state system: a chain in an open state (o) or a chain with a closed loop (c). At equilibrium, the interconversion $o \rightleftharpoons c$ obeys a detail balance condition: $\rho_c k_{c \rightarrow o} = \rho_o k_{o \rightarrow c}$ [6], where $\rho_{c/o}$ are the equilibrium densities and $k_{c \rightarrow o}, k_{o \rightarrow c}$ are the conversion rates. Loop closing and opening times, τ_c and τ_o , are inverses of the respective rates. Assuming a small interaction radius a between the chain ends, approximately $\rho_c = \frac{4\pi}{3} a^3 P_N(\vec{R} = 0)$, where $P_N(\vec{R})$ is the probability of finding the chain ends at a separation \vec{R} . But since possible number of open configurations far outweighs number of closed configurations i.e., $\rho_o \gg \rho_c$, we can approximate $\rho_o = 1 - \rho_c \simeq 1$ [6]. Further, τ_o is determined by the high energy barrier of the five bases long zipping strand, which is independent of the chain length N , and hence is a constant. Thus we arrive at $\tau_c \propto \rho_c^{-1}$, which we can compute from our model, as a function of chain length (N) and temperature, at zero force. In Fig.2 we plot $\tau_c/N^{3/2}$ versus inverse temperature, obtained from our model and compare it with the experimental data of Goddard et al [4]. The rationale for rescaling τ_c by $N^{3/2}$ is to partially nullify the strong N dependence in τ_c . Although $\rho_c \propto N^{-3/2}$ only for an FJC model (at large N) and not for an WLC model, nevertheless it turns out to be useful in approximately collapsing both our simulation data and Goddard et al's experimental data in a narrow range of N , near the room temperature.

Temperature dependence of extension : Change of extension with temperature has been shown to give interesting behavior for wild type ssDNA [19]. Fig.3 shows the analogous property resulting from our model of polydA, at fixed force, computed using transfer matrix method. In the absence of any transitions polymer extension is expected to decrease with temperature because of entropic elasticity, as it happens in rubber, for example. In case of wild type ssDNA hairpin loops can form which modifies the extension-temperature behavior in non-trivial ways. For polydA although loops cannot form in the absence of complementary bases, existence of two transitions (helix-coil and overstretching) makes it behave in an interesting way (see Fig3A), to the extent that weak, nonmonotonic, re-entrant behavior can be observed (Fig3B,C). These can be understood qualitatively. Two important points to remember here are, a) helix-coil transition can also be affected by raising temperature, and b) entropic fluctuations are enhanced at high temperatures which smoothens the force driven transitions (data not shown). Fig3A shows that at very low force like $5pN$, indeed extension weakly decreases with

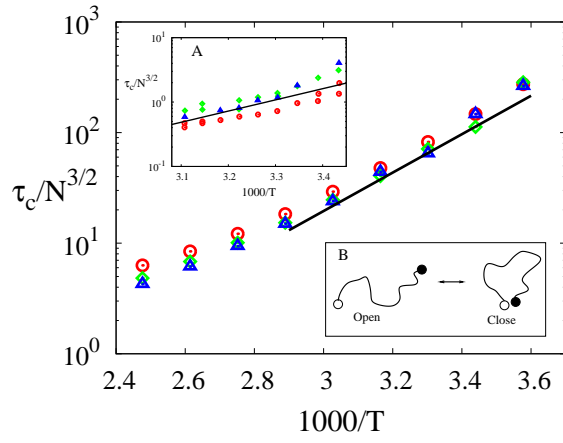


FIG. 2: (color online) Semi-log plots of the scaled loop closing time τ_c (of polydA) versus inverse temperature, computed for various chain lengths $N = 10$ (circle), 20 (diamond) and 30 (triangle), by Monte-Carlo simulation of our model. In a narrow window of temperature $1000/T \in [3.1, 3.45]$, the scaling of τ_c has been experimentally shown (inset-A) to be approximately, Arrhenius type (i.e., $\tau_c \propto \exp(\epsilon/k_B T)$) by Goddard et al [4]. Their data for chain lengths $N = 12$ (circle), 21 (diamond) and 30 (triangle), are plotted after rescaling with $N^{3/2}$). Both the Arrhenius aspect as well as the quantitative value of the exponent ϵ are closely reproduced by our model. The dashed lines in both the main plot and the inset have the same exponent $\epsilon = 13.4k_B T_o$, where $T_o = 298K$ is the room temperature. Inset-B schematically shows transition between open and closed chain conformations.

temperature but when force rises ($10pN$ onwards) as we move close to the helix-coil transition, it is easier to affect the transition by raising temperature and helix-coil transition leads to rise in the extension. But at forces just beyond the helix-coil transition ($30pN$ onwards but much below $100pN$) the ssDNA cannot access the overstretched bond lengths solely by means of thermal fluctuations and looses out to entropic elasticity showing decrease in extension with temperature. But as force approaches $100pN$ due to its vicinity to the overstretching transition the extension again increases first but eventually loose out to entropic elasticity at higher temperature, giving rise to reentrant behavior. Beyond the overstretching transition extension again decreases weakly with temperature due to entropic elasticity, weakly because at such high force not much entropy is left in the almost straight configuration.

In conclusion, we have proposed a new model for polydA, that incorporates two transitions and quantitatively reproduces both force-extension characteristics and loop closing statistics of such homopolymers. Our model also predicts interesting reentrant behavior in the temperature-extension diagram of polydA which can be verified experimentally.

We thank Dibyendu Das for useful comments.

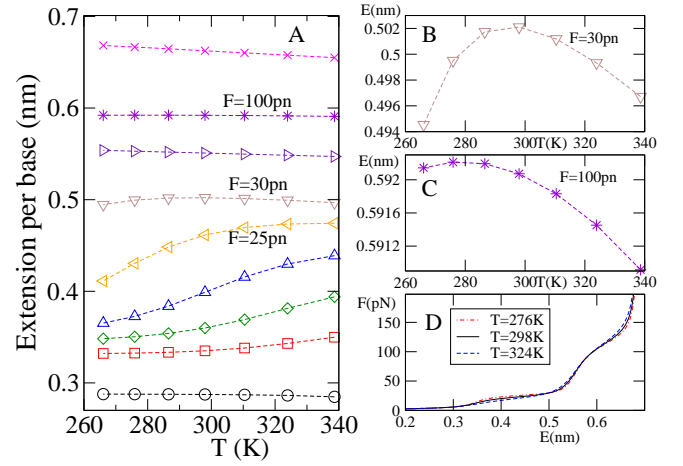


FIG. 3: (color online) Extension(E) per base (in nm) versus temperature, at a fixed force, plotted in (A) for different values of forces: 5, 10, 15, 20, 25, 30, 60, 100 and $150pN$, from bottom to top. These are obtained from our model using transfer matrix method. Effect of the first transition (near $25pN$) is reflected in the relatively large jump in the extension. Near the overstretching transition the jump is not so pronounced due to the relatively small plateau width in the corresponding FX diagram. The variation of extension with temperature changes qualitatively with force. The extension weakly decreases with temperature at low force, increases with temperature at intermediate force ($< 30pN$) and finally again decreases with temperature at very high force. The reentrant behavior i.e., extension initially increasing with temperature and later at higher temperature, decreasing with temperature, is zoomed in (B) and (C). The minute change in the FX curves with temperature is shown in (D). Dashed lines are guide for the eyes.

* Electronic address: asain@phy.iitb.ac.in

- [1] S. B. Smith, L. Finzi, C. Bustamante, Science, **258**, 1122 (1992).
- [2] C. Bustamante, J. Marko, E. Siggia, S. Smith, Science, **265**, 1599 (1994).
- [3] S. B. Smith, Y. Cui, C. Bustamante, Science, **271**, 795 (1996).
- [4] N.L. Goddard, G. Bonnet, O. Krichinsky and A. Libchaber, Phys. Rev. Lett. **85**, 2400 (2000).
- [5] D.P. Aalberts, J.M. Parman, N.L. Goddard, Biophys. J. **84**, 3212 (2003).
- [6] A. Sain, B.Y. Ha, H.K. Tsao, and J.Z.Y. Chen, Phys. Rev. E **69**, 061913 (2004). Phys. Rev. E **71**, 051902 (2005).
- [7] C. Ke, M. Humeniuk, H. S-Gracz, P. E. Marszalek Phys. Rev. Lett. **99**, 018302 (2007).
- [8] Wuen-shiu Chen et al. Phys. Rev. Lett. **105**, 218104 (2010).
- [9] B.H Zimm and J.K. Bragg, J. Chem. Phys. **28**, 1246 (1958); ibid **31**, 526 (1959).
- [10] Y. Seol, G. M. Skinner and K. Visscher Phys. Rev. Lett. **98**, 158103 (2007).
- [11] A. Buhot and A. Halperin, Phys. Rev. Lett. **84**, 10 (2000); Macromolecules **35**, 3238(2002).
- [12] M.N. Tamashiro and P. Pincus, Phys. Rev. E. **63**, 021909 (2001).

- [13] A. Ahsan, J. Rudnick and R. Bruinsma, *Biophys. J.* **74**, 132 (1998).
- [14] P. Cluzel et al., *Science* **271**, 792 (1996).
- [15] G. Mishra, D. Giri and S. Kumar, *Phys. Rev. E.* **79**, 031930 (2009).
- [16] S. Kumar and G. Mishra, *Soft Matter*, **7**, 4595 (2011).
- [17] Y. Seol, G. M. Skinner and K. Visscher *Phys. Rev. Lett.* **93**, 118102 (2004).
- [18] S.K. Ghosh, K. Singh and A. Sain, *Phys. Rev. E* **80**, 051904 (2009).
- [19] C. Danilowicz, C. H. Lee, V. W. Coljee, and M. Prentiss, *Phys. Rev. E.(R)* **75**, 030902 (2007).